

AD-A198 524

AUDITORY-ACOUSTIC BASIS OF CONSONANT PERCEPTION(U)  
CENTRAL INST FOR THE DEAF ST LOUIS MO J D MILLER  
30 OCT 87 AFOSR-TR-87-1764 AFOSR-86-0335

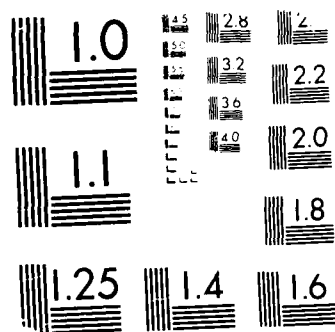
1/1

UNCLASSIFIED

F/G 25/4

ML





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

NONCLASSIFIED

AD-A190 524

## DOCUMENTATION PAGE

Form Approved  
OMB No. 0704-0188

UNCLASSIFIED		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY UNCLASSIFIED		3. DISTRIBUTION/AVAILABILITY OF REPORT UNCLASSIFIED	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE JAN 19 1988		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR-87-1764	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) D		7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research	
6a. NAME OF PERFORMING ORGANIZATION Central Institute for the Deaf		7b. ADDRESS (City, State, and ZIP Code) Bolling Air Force Base District of Columbia 20332	
6b. OFFICE SYMBOL (if applicable)		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-86-0335	
8a. NAME OF FUNDING / SPONSORING ORGANIZATION same as 7a AFOSR		10. SOURCE OF FUNDING NUMBERS	
8b. OFFICE SYMBOL (if applicable) NL		PROGRAM ELEMENT NO 61102F	PROJECT NO 2313
8c. ADDRESS (City, State, and ZIP Code) Bolling Air Force Base District of Columbia 20332		TASK NO A6	WORK UNIT ACCESSION NO
11. TITLE (Include Security Classification) Auditory-Acoustic Basis of Consonant Perception			
12. PERSONAL AUTHOR(S) James D. Miller			
13a. TYPE OF REPORT Annual Technical	13b. TIME COVERED FROM 9-30-86 TO 9-29-87	14. DATE OF REPORT (Year, Month, Day) 87-30-10	15. PAGE COUNT 12
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB-GROUP	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) Among the most interesting examples of the perception of complex sounds is that of the perception of consonants. Here, sequences of changing spectra induce the perception of phonetic entities in a manner that requires an understanding of the role of spectral trajectories, brief silences, the growth and decay of loudness, as well as language learning. An extensive study of the entire set of the consonant sounds of English is designed to elucidate, in quantitative detail, the sensory and perceptual processes whereby the acoustic waveform of speech is transformed by a series of processes leading to the perception of consonants as phonetic elements. Recordings of consonants as spoken by at least four male and four female talkers are to provide sample waveforms of each of the consonants of English in a variety of syllabic contexts, in common words, and in fluent phrases and sentences. These are being analyzed by a variety of digital techniques and the results are being interpreted in terms of a unifying theory of phonetic (Cont.)			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL JOHN F. TANGNEY		22b. TELEPHONE (Include Area Code) (202) 767-5021	22c. OFFICE SYMBOL NL

## 19. Abstract

perception -- the Auditory-Perceptual Theory. The method of synthesis is applied to determine those characteristics of the acoustic waveform essential to consonant perception and to evaluate the mathematical parameters in the Auditory-Perceptual Theory. Previous literature is being reanalyzed in terms of these new concepts. Among the experiments of interest are those on the categorical perception of speech sounds as exemplified by the pioneering experiments of Abramson and Lisker on the voiced-voiceless distinction in stops, and the experiments on cue integration (including silence) of Liberman and his associates at Haskins Laboratories. Also included is a significant effort in preparing slides, video tapes, and/or films that will illustrate the theoretical structures, both static and dynamic, in three-dimensional displays in both black and white and color. The overall goal of this research program is to extend work now underway on vowels and diphthongs (NIH Grant NS 21994-04) to include all of the phonetic elements of English. This is to provide a detailed account of the auditory-perceptual processes of phonetic perception by the human listener and, at the same time, provides a foundation for phonetically based automatic speech recognition which should be essentially independent of speaker and rate, with unlimited vocabulary in fluent speech.

AFOSR-TR. 87-17647-

AUDITORY-ACOUSTIC BASIS OF CONSONANT PERCEPTION

James D. Miller  
Central Institute for the Deaf  
818 S. Euclid Ave  
St. Louis, MO 63110

30 October 1987

Annual Technical Report for Period 30 September 1986 - 29 September 1987

Prepared for

Department of the Air Force  
Air Force Office of Scientific Research (AFSC)  
Bolling Air Force Base, DC 20332-6448

Accession For	
NTIS CRAWI	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Date	
Availability Codes	
Avail and/or	
Restrictions	
A-1	



Annual Technical Report  
Grant AFOSR 86-0335  
For the period 30 September 1986 - 29 September 1987

## I. SUMMARY

The first year of work on the auditory-perceptual basis of consonant perception has been marked not only by a variety of new and unique tools for the study of speech through computer programs and three-dimensional computer graphics but also by the discovery of new approaches to the acoustic and perceptual characterization of consonants. Additionally, measurements have been made of a variety of natural productions of consonants as well as of certain synthetic sounds that have played a key role in speech research over the past two decades.

The most significant achievement is the development of a promising algorithm for finding the perceptually significant features of burst-friction sounds which constitute an important fraction of speech and have resisted convenient analysis and succinct characterization. Using this algorithm, studies of the voiceless fricative consonants [s,sh,th,f,h] have been undertaken and are nearing completion. Similarly, studies of the bursts of the stop consonants [p,t,k,b,d,g] are nearing completion. Preliminary studies of the approximant consonants [l,r,w,j] have also been initiated and have provided preliminary results.

## II. RESEARCH OBJECTIVES

Among the most interesting examples of the perception of complex sounds is that of the perception of consonants. Here sequences of changing spectra induce the perception of phonetic entities in a manner that requires an understanding of the role of spectral trajectories, brief silences, the growth and decay of loudness, as well as language learning. An extensive study of the entire set of the consonant sounds of English is designed to elucidate, in quantitative detail, the sensory and perceptual processes whereby the acoustic waveform of speech is transformed by a series of processes leading to the perception of consonants as phonetic elements. Recordings of consonants as spoken by at least four male and four female talkers are to provide sample waveforms of each of the consonants of English in a variety of syllabic contexts, in common words, and in fluent phrases and sentences. These are being analyzed by a variety of digital techniques and the results are being interpreted in terms of a unifying theory of phonetic perception — the Auditory-Perceptual Theory. The method of synthesis is applied to determine those characteristics of the acoustic waveform essential to consonant perception and to evaluate the mathematical parameters in the Auditory-Perceptual Theory. Previous literature is being reanalyzed in terms of these new concepts. Among the experiments of interest are those on the categorical perception of speech sounds as exemplified by the pioneering

experiments of Abramson and Lisker on the voiced-voiceless distinction in stops, and the experiments on cue integration (including silence) of Liberman and his associates at Haskins Laboratories. Also included is a significant effort in preparing slides, video tapes, and/or films that will illustrate the theoretical structures, both static and dynamic, in three-dimensional displays in both black and white and color. The overall goal of this research program is to extend work now underway on vowels and diphthongs (NIH Grant NS 21994-04) to include all of the phonetic elements of English. This is to provide a detailed account of the auditory-perceptual processes of phonetic perception by the human listener and, at the same time, provides a foundation for phonetically based automatic speech recognition which should be essentially independent of speaker and rate, with unlimited vocabulary in fluent speech.

### III. STATUS OF THE RESEARCH

#### A. Location of burst-friction sounds in the Auditory-Perceptual Space

An important step in the analysis of stop consonants and fricatives in the Auditory-Perceptual Theory was the development of an algorithm to automatically determine the spectral prominences which characterize burst-friction sounds and to convert these into locations in the APS. This algorithm was developed by Allard Jongman. After a burst-friction sound is analyzed by means of LPC using a 24-ms full Hamming window moving in 1-ms steps, an FFT procedure extracts the formant information. The algorithm is then applied to the results of the FFT for every ms of burst-friction sound. First, the spectral peak with maximum amplitude below 6kHz is located and labeled P(max). Then moving from 60 to 6000 Hz, the first two peaks within 10 dB of P(max) are picked as the Burst-Friction Sensory Formants, BF2 and BF3. Furthermore, in those cases where BF2 has been picked and BF3 is separated from BF2 by 2500 Hz or more, the frequency value for BF2 is also used as that for BF3. Finally, if there are no peaks within 10 dB of P(max), the frequency value of the maximum peak is used for both BF2 and BF3.

This algorithm has been implemented in software by Steven J. Sadoff and enables us to automatically extract spectral information characteristic of burst friction sounds such as stop consonants and fricatives from the output of the FFT procedure.

#### B. Stop consonants

##### 1. Place of articulation in voiceless stop consonants

Two male and two female speakers were recorded, producing two repetitions each of real CVC words, with the three voiceless stop consonants ([p,t,k]) in initial position and followed by the vowels ([i,I,ɛ,a,ʊ,u]). In order to determine the locations of burst onsets in the algorithm described under III.A was applied to only the first millisecond of the stop-burst. The frequency values of BF2 and BF3 that were extracted by the algorithm were then converted into coordinates of the auditory-perceptual space (APS) using the following equations:

$$x = \log(\text{BF3}/\text{BF2})$$

$$y = \log(SR/SR)$$

$$z = \log(BF2/SR)$$

$$SR = 160(GMTFO/168)^{1/3}$$

where GMTFO is the estimated geometric mean of the current speaker's FO.

Since burst-friction sounds do not have a first formant, BF1 is arbitrarily set equal to SR. In these cases, y equals 0, and, therefore, burst-friction sounds are located in the xz-plane of APS.

The x and z coordinates associated with each burst onset were then plotted in APS. Bilabial, alveolar, and velar burst-onset target zones were drawn in an attempt to minimize overlap. This method of analysis enabled us to identify place of articulation in voiceless stops with 89% accuracy.

The algorithm and results of this study were presented by A. Jongman (1987) at the Fall meeting of the Acoustical Society of America in Miami, Florida.

## 2. The voicing distinction in stop consonants

In order to analyze the voicing distinction in terms of APT, the original Abramson and Lisker (1970) synthetic voice-onset-time (VOT) continua were used. The three continua, one for each place of articulation ([ba-pa], [da-ta], [ga-ka]), consisted of 18 members each (0-85 ms VOT in 5-ms steps). The burst-friction components of these synthetic stimuli were analyzed using the algorithm described in III.A, and for each millisecond of the burst-friction segment, x and z coordinates were plotted, resulting in a burst-friction sensory path in APS.

For the glottal-source component (corresponding to the vowel [a]), the first three sensory formants (SF1, SF2, and SF3) were extracted and converted into APS coordinates using the following equations:

$$x = \log(SF3/SF2)$$

$$y = \log(SF1/SR)$$

$$z = \log(SF2/SF1)$$

For each millisecond of the glottal-source segment, x, y, and z coordinates were plotted, resulting in a glottal-source sensory path in APS. The sensory-perceptual transformation then serves to integrate burst-friction and glottal-source components into a unitary response called a perceptual path, using a spring mass model. Perceptual paths were then plotted in APS for each continuum member. Observation of these paths revealed that all VOT-continuum members entered the appropriate stop target zones (labial, alveola, and velar). These target zones were extrapolated from the burst-onset target zones that were described in III.B. For example, all members of the [ba-pa] continuum first entered the bilabial stop target zone. In addition, the vocalic part always entered the target zone for [a].



Given that all stimuli entered the appropriate target zones in terms of their place of articulation, the next issue was to determine how voiced and voiceless stops are distinguished in the auditory-perceptual theory (APT). In this regard, it is important to note that for English stops the voicing distinction is a distinction between voiceless unaspirated (e.g., [p]) and voiceless aspirated (e.g., [p<sup>h</sup>]) stops. That is, English listeners will perceive voiceless unaspirated stops as voiced, and voiceless aspirated stops as voiceless.

We hypothesized that the activation of the aspirated [h]-target zone (described in III.C), or lack thereof, is one way of distinguishing voiced stops from their voiceless counterparts. Results can be summarized as follows:

- short VOT stimuli did not enter the [h]-target zone; instead, they entered the appropriate stop target zone and then enter the [a] target zone.
- long VOT stimuli entered appropriate stop target zone and then entered and reached the center of the [h]-target zone before entering the [a] target zone.
- VOT boundary stimuli ([b/p]-VOT=25 ms, [d/t]-VOT=35 ms, [g/k]-VOT=45 ms) approached the border of, but did not enter, the [h]-target zone.

These preliminary results suggest that the concept of the sensory-perceptual transformation, the stop target zones, and the [h]-target zone enable us to distinguish voiced and voiceless English stop consonants in a way consistent with experimental data on categorical perception.

The results of this study were reported by J.D. Miller and A. Jongman (1987) at the Fall meeting of the Acoustical Society of America in Miami, Florida.

#### C. Fricatives

Three male and three female speakers were recorded, producing one token each of CV syllables with [f, θ, s, ʃ] in initial position, followed by each of the vowels [i, u, a]. For [h], two male and two female speakers were recorded, producing two repetitions each of [hVd] words, where V is each of the 10 simple vowels of American English [i, I, ε, e, a, ʌ, ɔ, ʊ, ɜ, u]. The algorithm described in 2.1.1 was applied to each of the burst-friction segments, and the geometric means of BF2 and BF3 over the entire burst-friction segment were converted into x and z coordinates. These coordinates were plotted in APS, and target zones were drawn.

In this way, [s] was distinguished from [ʃ] with 100% accuracy. However, [f] and [θ] could not be differentiated, a notorious problem in the speech literature, and the [h]-target zone showed considerable overlap with those of [f] and [θ].

#### D. Approximants [l, r, w, j]

We have measured l's and w's in the syllables: wheel, will, well, wall, la, lae, lull, wool, and woo as spoken by two male and two female talkers. Additionally, we have measured w's, r's, and j's in the sentence "where were you a year ago." Based on these observations and on data taken from literature the target zones for these sounds are being revised.

#### E. Software Development

We have made considerable progress in developing software for the implementation of the theory on computers, using both the Evans and Sutherland three-dimensional graphics terminal and regular two-dimensional terminals. Below we report the work of the last two years. The first year sponsored by the NIH Grant (NS 21994-04) and the second year jointly supported by the NIH Grant and the AFOSR Grant that is the subject of this report.

The Evans and Sutherland PS300, a high speed, high resolution color graphics system, and its VAX-VMS host system are used to display, manipulate, and analyze objects in the three-dimensional auditory-perceptual space. In the majority of cases, software used in this research effort has been specially developed for these rather specialized applications. The programs MWVNET, DISPLAY and SLICER are the three most used application programs and are described below.

MWVNET is a PS300 function network that allows the user to examine an object and manipulate it in four different coordinate systems: world, model, part and view. The program implements keyboard commands for the choice of coordinate systems and rotary dial input for scaling, translation and rotation of the displayed objects. MWVNET also forms the framework for most of the other application programs written for speech perception studies on the PS300.

DISPLAY is an application program whose primary function is to provide a user interface for the display and manipulation of objects defined in PS300 code. It has facilities for highlighting, blinking, coloring and hiding objects. It also provides an interface for operations involving the host system such as running command files and the downloading of object data files from memory. Several important features have been recently implemented that expand the use of DISPLAY as a research tool. The program now has the capability to identify and separate the burst-friction and glottal-source sections of a sensory path into individually defined and manipulable objects. In addition, the user may now "track" along a sensory or perceptual path with a cursor and obtain the x, y, z and x', y', z' coordinates of any point along the path as well as an average value for points in a user-determined subsection. This feature is invaluable in the choice of a target point for a particular section and the subsequent construction of a target zone from collections of such points. Hard copy plots of displayed data may now be generated with a six-pen plotter or with an Apple LaserWriter. Such plots may be used for journal quality reproductions of auditory-perceptual data.

SLICER is a program used in the construction of wireframe target zones that surround point data in the three-dimensional auditory-perceptual space. It displays successive slices of target data allowing the user to draw delimiting outlines around the two-dimensional slice of a target zone. This is a

computerized version of the method of serial sections that has been usefully applied in microanatomical studies for many years. The vector lists which comprise the slice traces are converted to raster scans by the program CNTSYB. The rasterized data are then contoured into a three-dimensional wireframe model that represents the target zone by the commercial package SYBYL. The PS300 code which represents the target zone is then compressed by the program VCOMPRESS which reduces the amount of storage required for the target zone by as much as 80 percent.

Additionally, since a great part of the work preliminary to plotting on the Evans and Sutherland is done using two-dimensional graphics terminals, we have developed a set of software packages, which allow us to digitize and edit waveforms as well as produce plots of all the variables involved in the auditory-perceptual theory on such terminals. First, in order to simplify and standardize the writing of software which utilizes graphics, we have developed a set of 2-dimensional graphics subroutines and compiled these into a library which we call PLOT10. This library provides a functionally complete graphics interface to any device that can emulate the Tektronix 4010 series of terminals. This library has enabled us to develop many applications that can display graphics. It has allowed researchers whose only familiarity with computers is FORTRAN to develop graphics software without involving them in the details of sending escape sequences and cryptic address coordinates. To handle the various peculiarities of different PLOT10 emulations at run-time (as opposed to compile or link time), this graphics package utilizes a system-wide text file that describes the individual characteristics of the particular terminal type being used. In this file we store items such as terminal resolution and escape sequences for entering and exiting graphics mode. This frees the programmer from dealing with the intricacies of each particular terminal, providing some degree of device independence. This package works on all of the terminals that we have access to including DEC VT240's, MicroTerm Ergo-301's, Graphon GQ-140's, and HP2623's. Hardcopy can either be obtained by screen dumps from any of our HP2623s or we can direct the graphics package to use our LN03 laser printer for publication-quality output. These routines were meant to be called from FORTRAN, but if the proper calling conventions are maintained, they may be called from any other language.

Two other important graphics routines have been developed. These are FMPTL and VAK. FMPTL allows the user to plot the values of the sensory variables SR, SF1L, SF1H, SF2, SF3, BF2, and BF3 as a function of time or as a function of distance traveled on the corresponding perceptual path in the APS. The user may choose either a logarithmic or linear frequency scale. BF2 and BF3 are clearly distinguished from SF2 and SF3 by the use of x's rather than dots. Options to plot F0 are available and options to plot F0 modulations are planned. FMPTL also allows the user to simultaneously plot the perceptual variables PR, PF1L, PF1H, PF2, and PF3 against time or distance. Once again one may choose a linear or log frequency scale. These programs allow the user to directly view these formant tracks and compare them to what is seen in the spectrogram, what is heard, and what is observed in the APS. A variety of cursor options are planned. The graphics package VAK is oriented to the auditory-perceptual space and the search for segmentation rules. The user may plot APS coordinates (x, y, z) or slab coordinates (x', y', z'). Either sensory or perceptual values may be

selected and these may be plotted against time or distance traveled. Cursor options allow one to determine the exact values of the plotted functions at any point along the curve. Additionally one may plot distance in the APS against elapsed time or the magnitude of velocity and acceleration of the perceptual pointer in APS against time or distance, with magnification of the variables and cursor measures as options. Similarly, an index of path curvature can be plotted against time and distance. Another set of VAK options includes plotting the signed velocity of the perceptual pointer in each of the dimensions  $x$ ,  $y$ ,  $z$ ,  $x'$ ,  $y'$ , or  $z'$  against either time or distance. These routines now allow us to quickly evaluate a variety of variables implicated as contributing to segmentation. In addition, a third routine, MULTPA, plots sensory and perceptual paths on a two-dimensional screen, along with as many target zones as the user specifies. Options include front view of the vowel space or sideview, line vs discrete symbols for each data point, and dumping to the laser printer for publication-quality output. Future work will add intensity and pitch information to the battery of plots.

We also developed our own digitization and waveform-editing package named SINS. SINS is an interactive graphical editor designed to work with a DigiSound-16 system connected to a MicroVAX II using a SAP interface along with a DRV11-WA. SINS is an acronym for Speech IN the auditory perceptual Space. It is used for controlling analog-to-digital (A/D) and digital-to-analog (D/A) operations, as well as performing simple editing and windowing operations upon sampled waveforms. Currently we are using SINS to digitize our audio tapes recorded on our JVC VCR in our anechoic chamber. SINS is capable of reading and writing many different file formats including files which are compatible with ILS, the commercially available signal processing package that we are currently using. We have tested SINS with as many as 16 users logged onto the system at once, indicating that it is feasible to do many real-time operations on a multi-user computer running the VMS operating system. The software was written in a modular fashion so that SINS never accesses the Digisound directly. All I/O for the DigiSound-16 system is performed through the Digisound-16 library which we have developed. There are only three SINS commands that call routines from the DigiSound-16 library: play, record, and setting the sampling rate. To enable this package to work with a different D/A-A/D system would simply require a rewrite of these three routines. SINS will provide a graphical interface, if the user is using a Tektronix 4010 compatible terminal. All graphics operations are performed using the PLOT10 library, allowing this software to be used on any type of terminal supported by the PLOT10 library. To enable the graphics to work with a different type of terminal, would simply require modifying the PLOT10 package. This should allow this software to be ported in the future to other terminal types. All other screen I/O (user input and prompting) is performed using the standard DEC Screen Management routines (SMGS Run Time Library).

Finally, we developed a program to assist the user in editing a file which contains a list of the formants (an FMT file). The FMT file is obtained by running the program GETFIF on an analysis file which has undergone an API and a SGM (Analysis commands of the Interactive Laboratory System package). The program, named INTER, is used mainly to correct inaccuracies in our formant tracking. Of the many options, geometric interpolation, and linear

interpolation are used quite frequently. Additionally, INTER can calculate the values for F1L and F1H for segments containing a voice bar. Also, columns of formant values can be copied into other columns, since a common problem with our current formant tracking is the mislabeling of formants (i.e. when F2 and F1 merge, the values placed in F2 really are values for F3). This is necessary since we do not have access to an editor that has select and paste operations on columns.

We have also implemented the Klatt synthesis program on our MicroVAX II. This program has been modified in several ways by adding subroutines that enhance the front end. We now have options for different input glottal waveforms and output directly to an ILS file. We can now also use a digitizer pad with the front view of the vowel slab to enter  $x'$ ,  $y'$  coordinates with a pre-set  $z'$ . These values are then automatically converted to formant values and bandwidths which are used as parameters for synthesis. A separate program has been written which allows batch synthesis overnight of great numbers of stimuli without requiring the presence of the experimenter. This capability will now be used to precisely define the borders of the target zones.

These packages, all of them developed in the last two years, provide an excellent environment for carrying out our research. We now plan to enhance the software, as we keep developing the auditory-perceptual theory, so that the end result will be a hands-off processing of the acoustical signal of speech. All of these programs are necessary to enable us to conduct our basic research on human speech perception.

#### IV PUBLICATIONS AND MANUSCRIPTS BY PARTICIPANTS

- Bowe, C.A., Green, L. and Miller, J.D. (1987). "Differential acquisition of discriminated autoshaping as a function of stimulus qualities and location," Animal Learning & Behavior, 15 (3), pp. 285-292.
- Bowe, C.A., Miller, J.D., and Green, L. (1987). "Qualities and locations of stimuli and responses affecting discrimination learning of chinchillas (Chinchilla laniger) and pigeons (Columba livia)," J. Comp Psych., 101 No.2, pp.132-138.
- Chang, H.M. (1987). "SWIS: See what I say, a speaker-independent word recognition system by phoneme-oriented mapping on aphonetically-encoded auditory-perceptual speech map," Ph.D. dissertation, Washington University, St. Louis, MO.
- Fourakis, M. (1986). "An acoustic study of the effects of tempo and stress on segmental intervals in modern Greek," Phonetica, 43, pp. 172-188.
- Fourakis, M. (1986). "A timing model for word initial CV-syllables in Modern Greek," J. Acoust. Soc. Am., 79, pp. 1982-1986.
- Miller, J.D. (1987). "Auditory-perceptual processing of speech waveforms," In: W.A. Yost, and C.S. Watson (Eds.) Auditory Processing of Complex, Lawrence Erlbaum: Hillside, N.J., pp.257-266.
- Osberger, M.J., Johnson, D.L., and Miller, J.D. (1987). "Use of connected discourse tracking to train functional speech skills," Ear & Hearing, 8, pp. 31-36.
- Skinner, M.W., Miller, J.D., DeFilippo, C.L., Dawson, J.K., and Popelka, G.R. (1986). "Word identification by listeners with sensorineural hearing loss using four amplification systems," In: M.J. Collins, T.J. Gattke, and L.A. Harker (Eds.) Sensorineural Hearing Loss, Iowa City, Iowa: University of Iowa, pp. 305-325
- Weisenberger, J.M. and Miller, J.D. (1987). "The role of tactile aids in providing information about acoustic stimuli," J. Acoust. Soc. Am., 82 (3), pp. 906-916.
- Boettcher, F. and Miller, J.D. (in press). "The formation of learning sets by chinchillas in an auditory discrimination task." J. Auditory Res.
- Jongman, A. and Miller, J.D. (in preparation). "Location of stop-bursts in the auditory-perceptual space."
- Miller, J.D. and Fourakis, M.S. (in press). "Speech Perception," International Encyclopedia of Communication.

- Miller, J.D. "Auditory-perceptual theory of phonetic recognition," (submitted to J. of Phonetics).
- Miller, J.D., "Auditory-perceptual interpretation of the vowel," (submitted to J. Acoust. Soc. Am.).
- Miller, J.D., Fourakis, M.S., Sadoff, S.J., Veksler, M.R., Jongman, A., Kramer, F.E. and Hawks, J.W. (in preparation). "Methods for the study of speech spectra and their dynamics."
- Miller, J.D., Sadoff, S.J., and Chang, H.M. (in preparation). "The sensory-perceptual transformation in speech perception."
- Fourakis, M.S. and Miller, J.D. (in preparation). "Measurements of vowels in isolation and in sonorant context."
- Fourakis, M. and G.K. Iverson (in press). "More on second-language timing patterns," Language Learning.

The following abstracts submitted by members of the Speech Perception Laboratory have been accepted for or published by various conferences.

- Chang, H.M. (1987). "Automatic detection for diphthongs," J. Acous. Soc. Am., 82(S1), S37(A).
- Jongman, A. (1987). "Locations of the burst onsets of stops auditory perceptual space." J. Acous. Soc. Am., 82(S1), S82(A).
- Miller, J.D. and Jongman, A. (1987). "Auditory-perceptual approach to stop consonants." J. Acous. Soc. Am., 82(S1), S82(A).
- Miller, J.D. (1987). "Classification of vowel productions by means of perceptual target zones: A response to Ladefoged and Studdert Kennedy." J. Acous. Soc. Am., 82(S1), S82(A).
- Miller, J.D. (1987). "Auditory-perceptual interpretation of the vowel," J. Acoust. Soc. Am., 81(S1), S16(A).
- Jongman, A. and Fourakis, M.S. (1987). "A cross-language study of vowel spaces and interference," J. Acoust Soc. Am., 81(S1), S66(A).
- Fourakis, M.S. and Miller, J.D. (1987). "Measurements of vowels in isolation and in sonorant context," J. Acoust. Soc. Am., 81(S1), S17(A).

## V PARTICIPATING PROFESSIONALS

### James D. Miller

University of Wisconsin, Madison, WI	B.S.	1951	Psychology
Indiana University, Bloomington, IN	M.A.	1953	Psychology
Indiana University, Bloomington, IN	Ph.D.	1957	Psychology

Dissertation title: "On the relationship between temporary hearing loss and masking."

### Marios S. Fourakis

Wabash College, Crawfordsville, IN	B.A.	1973	Classical Greek Lit.
Indiana University, Bloomington, IN	M.A.	1979	Classical Studies
Indiana University, Bloomington, IN	M.A.	1980	Linguistics
Indiana University, Bloomington, IN	Ph.D.	1983	Linguistics

Dissertation title: "An acoustic study of temporal programming in speech production."

### Allard Jongman

University of Amsterdam, The Netherlands	B.A.	1980	Slavics
University of Amsterdam, The Netherlands	M.A.	1982	Linguistics
Brown University, Providence, RI	Ph.D.	1986	Linguistics

Dissertation title: "Naturalness in phonetics: A study of context-dependency."

### Joan A. Sereno

Northern Illinois University, De Kalb, IL	B.S.	1982	Psychology
Northern Illinois University, De Kalb, IL	B.A.	1982	Philosophy
Brown University, Providence, RI	M.A.	1986	Linguistics
Brown University, Providence, RI	Ph.D.	1987	Linguistics

Dissertation title: "Graphemic, associative, and syntactic priming effects at a brief stimulus onset asynchrony in naming and lexical decision."

### Frank E. Kramer

University of Houston, Houston, TX	B.S.	1972	Geology
Washington University, St. Louis, MO	M.S.	1987	Geochemistry

Thesis title: "The distribution of Krypton in an anorthite-diopside-water mixture at five kilobars pressure."



Steven J. Sadoff

Washington University, St. Louis, MO	B.S.	1985	Computer Science
Washington University, St. Louis, MO	M.S.	1987	Electrical Eng.

Lynn W. Shields

University of Tennessee, Knoxville, TN	M.A.	1976	Speech Pathology
--	------	------	------------------

Mark R. Veksler

Washington University, St. Louis, MO	B.S.	1986	Systems Science&Math
--------------------------------------	------	------	----------------------

VI PRESENTATIONS

Miller, J.D. and Hawks, J.W., Listener-talker interaction: Is there an "autophonetic" effect?, 113th meeting of the Acoustical Society of America, Indianapolis, IN., May 11-15, 1987.

Fourakis, M.S. and Miller, J.D., Measurements of vowels in isolation and in sonorant context., 113th meeting of the Acoustical Society of America, Indianapolis, IN., May 11-15, 1987.

Jongman, A. and Fourakis, M.S., A cross-language study of vowel spaces and interference., 113th meeting of the Acoustical Society of America, Indianapolis, IN., May 11-15, 1987.

Miller, J.D., Auditory-perceptual interpretation of the vowel., 113th meeting of the Acoustical Society of America, Indianapolis, IN., May 11-15, 1987.

Fourakis, M.S., Current speech research at Central Institute for the Deaf., Haskins Laboratory, New Haven, CN., August 25, 1987.

Miller, J.D., "The auditory-perceptual theory of phonetic recognition," seminar presented at Central Institute for the Deaf, October, 16, 1986.

Miller, J.D., "Auditory-perceptual theory of phonetic recognition," invited lecture presented at AFOSR, Bolling AFB, April, 1986.

VII PATENT DISCLOSURES

Patent file numbers 5508A and 5508B	- Speech Processing Apparatus And Methods
Patent file number 5510	- Burst-Friction Formants Identification

END

DATE

FILMED

4-88

DTIC